

Package: domainator (via r-universe)

September 16, 2024

Title Classify Domains using Webservices and Lists

Version 0.0.0.9000

Description Wraps various APIs and provides lists to classify domains.
This includes...

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

URL <https://github.com/schochastics/domainator>

BugReports <https://github.com/schochastics/domainator/issues>

Imports base64enc, utils, httr2, rlang

Depends R (>= 2.10)

LazyData true

Repository <https://schochastics.r-universe.dev>

RemoteUrl <https://github.com/schochastics/domainator>

RemoteRef HEAD

RemoteSha 045a83a7d31c875d7840bfcfb843200ee94cfad1

Contents

domain_list	2
get_categories_shrinker	2
get_url_category_klazify	3
get_url_category_linkpreview	3
get_url_category_shrinker	4
get_url_category_whoisxml	5
news_types	5
online_news	6
us_news	6
Index	7

domain_list	<i>Domain list classification of domains into news,portals, search, and social media</i>
-------------	--

Description

Domain list classification of domains into news,portals, search, and social media

Usage

```
domain_list
```

Format

An object of class `data.frame` with 663 rows and 2 columns.

References

Stier, S., Mangold, F., Scharnow, M., & Breuer, J. (2022). Post Post-Broadcast Democracy? News Exposure in the Age of Online Intermediaries. *American Political Science Review*, 116(2), 768-774.

get_categories_shrinker	<i>List All Web Shrinker Categories</i>
-------------------------	---

Description

Returns all of the possible categories from `webshrinker.com` that URLs, hostnames, and IP addresses can be associated with.

Usage

```
get_categories_shrinker(taxonomy = "iabv1")
```

Arguments

`taxonomy` character. Which category taxonomy to use, either "iabv1" or "webshrinker"

Value

list of categories

Examples

```
## Not run:  
get_categories_shrinker(taxonomy = "iabv1")  
  
## End(Not run)
```

```
get_url_category_klazify  
    Klazify Category Lookup
```

Description

uses klazify.com to get the categories associated with the given URL.

Usage

```
get_url_category_klazify(url, raw = FALSE)
```

Arguments

<code>url</code>	character. URL to be categorized (needs https://www.)
<code>raw</code>	boolean. If TRUE returns the raw API response which contains more information than only the categories.

Value

data.frame with the inferred category and confidants score

Examples

```
## Not run:  
get_url_category_klazify("https://www.kicker.de")  
  
## End(Not run)
```

```
get_url_category_linkpreview  
    linkpreviewer
```

Description

use linkpreview to get the description of a URL.

Usage

```
get_url_category_linkpreview(url)
```

Arguments

url character. URL to be categorized

Value

data.frame with the title and description of a website

Examples

```
## Not run:  
get_url_category_linkpreview("https://www.google.com")  
  
## End(Not run)
```

get_url_category_shrinker
Webshrinker Category Lookup

Description

uses webshrinker.com to get the categories associated with the given URL, hostname, or IP address

Usage

```
get_url_category_shrinker(url, taxonomy = "iabv1")
```

Arguments

url character. URL to be categorized (does not need https://www.)
taxonomy character. Which category taxonomy to use, either "iabv1" or "webshrinker"

Value

data.frame with the inferred category and confidants score

Examples

```
## Not run:  
get_url_category_shrinker("kicker.de")  
  
## End(Not run)
```

```
get_url_category_whoisxml
```

Whoisxml Category Lookup

Description

uses whoisxmlapi to get the categories associated with the given URL

Usage

```
get_url_category_whoisxml(url)
```

Arguments

url character. URL to be categorized (doesnt need https://www.)

Value

data.frame with the inferred category and confidents score

Examples

```
## Not run:  
get_url_category_whoisxml("kicker.de")  
  
## End(Not run)
```

```
news_types                    News Types
```

Description

Classification of domains into different news types

Usage

```
news_types
```

Format

An object of class data.frame with 690 rows and 2 columns.

References

Stier, S., Mangold, F., Scharnow, M., & Breuer, J. (2022). Post Post-Broadcast Democracy? News Exposure in the Age of Online Intermediaries. *American Political Science Review*, 116(2), 768-774.

online_news	<i>Domain list of online news</i>
-------------	-----------------------------------

Description

Domain list of online news

Usage

online_news

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 542 rows and 2 columns.

References

Sebastian Stier, Nora Kirkizh, Caterina Froio, Ralph Schroeder. Populist attitudes and selective exposure to online news <https://osf.io/5pe27>

us_news	<i>Domain list of us news contains a list of and information on 5,397 US news domains. The domains overwhelmingly represent US-based organizations, but since the list was partly sourced from browsing data, it also includes some international domains visited by US study participants.</i>
---------	---

Description

Domain list of us news contains a list of and information on 5,397 US news domains. The domains overwhelmingly represent US-based organizations, but since the list was partly sourced from browsing data, it also includes some international domains visited by US study participants.

Usage

us_news

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 5397 rows and 3 columns.

References

Clemm von Hohenberg, B., Menchen-Trevino, E., Casas, A., Wojcieszak, M. (2021). A list of over 5000 US news domains and their social media accounts. <https://doi.org/10.5281/zenodo.7651047>

Index

* datasets

domain_list, [2](#)

news_types, [5](#)

online_news, [6](#)

us_news, [6](#)

domain_list, [2](#)

get_categories_shrinker, [2](#)

get_url_category_klazify, [3](#)

get_url_category_linkpreview, [3](#)

get_url_category_shrinker, [4](#)

get_url_category_whoisxml, [5](#)

news_types, [5](#)

online_news, [6](#)

us_news, [6](#)